

# La représentation des sens modaux dans trois langues romanes : le français, l'italien et le roumain. Du qualitatif au quantitatif et retour<sup>1</sup>

Claudia Ricci, Corinne Rossari (Université de Neuchâtel) et Elena Siminiciuc (Université d'Oxford)

## 1. Introduction

Pourquoi s'intéresser aux sens modaux dans une perspective comparative entre langues romanes ? La perspective inter-langue prenant en compte des langues de même souche fait ressortir des variations a priori inattendues concernant les sens modaux. Les mêmes sens modaux se stabilisent sur des morphèmes différents d'une langue à l'autre, alors que le sens lexical (celui qui correspond à la signification dite première ou littérale) est porté par la même forme. Notre étude prendra appui sur le futur. Le futur italien a des sens modaux que le futur français, qui partage la même morphologie, ne véhicule pas dans l'état actuel de la langue. Un de ces sens est le sens concessif, qui est transmis par *peut-être* en français contemporain.

- (1) **Sarò** italiana, ma non mi piace la pizza
- (2) Je suis **peut-être** italienne, mais je n'aime pas la pizza

En revanche, en roumain ce même sens modal peut également être transmis par une forme de futur (le présomptif) qui n'a pas la même morphologie que le futur italien<sup>2</sup>.

- (3) **Oi fi** eu italiana, dar nu-mi place pizza.

Par ailleurs, le roumain et l'italien, tout comme le français, peuvent utiliser l'adverbe correspondant à *peut-être* (*poate / forse*) pour transmettre ce sens.

- (4) *Sono forse Italiano per l'anagrafe perché ahimè sono nato lì, ma appena ho avuto l'opportunità, me ne sono andato il più lontano possibile.*

([www.australianboardcommunity.com](http://www.australianboardcommunity.com))

*Je suis peut-être italien pour le registre de l'état civil car hélas je suis né là-bas, mais dès que j'en ai eu l'opportunité, je m'en suis allé le plus loin possible.*

---

<sup>1</sup> Cette étude s'inscrit dans le cadre du projet financé par le FNS intitulé : « La représentation du sens modal et de ses tendances évolutives dans deux langues romanes : le français et l'italien » (projet no : 100012\_159458), ainsi que, pour le roumain, dans le cadre du projet : « Formes grammaticales et lexicales de la modalité en roumain » (projet no : P2FRP1\_161718). Nous remercions Annalena Hütsch, Maryana Kuzemko, Lamei Li, Margot Salsmann, Tobias von Waldkirch et Dennis Wandel pour leur collaboration au décompte des formes.

<sup>2</sup> Nous nous référons à l'ouvrage de Popescu (2013 : 20-23) pour la description du paradigme du présomptif. Ce paradigme est composé d'un ensemble de constructions périphrastiques, issues du paradigme du futur, plus ou moins grammaticalisées, qui connaît du point de vue temporel le présent (présomptif présent) et le passé (présomptif passé). Le paradigme du présent comprend deux sous-types morphologiques: (i) le type constitué de l'auxiliaire, par ex. 3ème pers. *o* (forme populaire érodée de *vouloir*) + *fi* (forme courte de l'infinitif du verbe *être*) + le gérondif du verbe lexical, par ex. *cântând* (chantant) : *o fi cântând* ; (ii) un paradigme formé au moyen d'un autre type de futur toujours composé de la forme populaire érodée du verbe *vouloir* : *o* + l'infinitif court du verbe lexical (*cânta*) : *o cânta*. Le paradigme du passé (présomptif passé) comprend un seul type de forme : l'auxiliaire *o* + l'infinitif du verbe *être* : *fi* + le participe passé du verbe lexical (*cântat*) : *o fi cântat*.

- (5) **poate** sunt eu roman, da' nu-s si poet  
*je suis peut-être roumain, mais je ne suis pas poète*  
(<https://mocioi.wordpress.com>)

Le fait que les mêmes formes développent des sens modaux différents d'une langue à l'autre interroge la façon dont ces sens se répartissent dans le lexique d'une langue et les motifs en fonction desquels ils se greffent sur certaines expressions ou marques grammaticales au détriment d'autres.

Après avoir adopté une approche qualitative pour représenter les sens modaux du futur dans Rossari *et al.* (à paraître), dans cet article nous intégrons à cette approche une démarche quantitative pour voir, au moyen d'un corpus constitué de micro-genres discursifs semblables d'une langue à l'autre, dans quelles proportions ces sens s'actualisent sur certaines formes au détriment d'autres au sein du lexique d'une langue donnée. Notre propos sera d'identifier des facteurs qui facilitent le développement de sens modaux. Notre recherche porte sur deux sens modaux : le sens épistémique et le sens concessif, qui peuvent être véhiculés par le futur ainsi que par des formes lexicales (pour une discussion du sens épistémique du futur, nous renvoyons à Dendale 2001 et à Morency 2010 pour le français et à Bertinetto 1991 pour l'italien. Pour une description du futur concessif italien, voir entre autres Berretta 1997 et Baranzini à paraître).

Pour le sens épistémique, en plus du futur, nous prenons en compte dans les trois langues les formes lexicales suivantes, susceptibles de véhiculer des valeurs approchantes : *devoir* (au sens épistémique), *peut-être*, *probablement*, *sûrement*, *sans doute*.

Pour le sens concessif, que le futur italien et le présomptif roumain peuvent véhiculer, nous prendrons en compte *peut-être* (2) en français, *forse* (4) en italien et *poate* (5) en roumain.

Avant d'analyser la fréquence de ces formes dans notre corpus, nous présenterons le cadre théorique que nous avons conçu pour représenter les différents sens modaux qu'elles peuvent véhiculer. Ce cadre nous sert de base pour identifier les usages des formes que nous considérons comme vectrices de sens modaux.

## 2. La représentation des sens modaux

Les formes qui véhiculent des sens modaux peuvent agir à différents niveaux concernant l'interprétation d'un énoncé.

Elles peuvent agir au niveau du contenu propositionnel (cp), soit parce qu'elles sont sémantiquement constitutives de ce dernier, comme la marque grammaticale du futur temporel, soit parce qu'elles sont sémantiquement constitutives du composant attitudinal qui porte sur ce contenu, comme *devoir* épistémique ou *peut-être*. Dans ce dernier cas, elles échappent aux conditions de vérité.

- (6) a. Paul est **peut-être** malade, il n'est pas venu au bureau / Paul **doit** être malade, il n'est pas venu au bureau  
b. Non, il est en pleine forme, je l'ai vu à la salle de gym.

Comme le préconisent les tests classiques, l'adverbe *peut-être* et le verbe épistémique *devoir* restent en dehors de la portée de la réfutation faite par l'interlocuteur. Ces formes agissent sur le contenu, mais ne sont pas sémantiquement constitutives de ce dernier.

Quand le futur est sémantiquement constitutif du cp, il est réfutable et est forcément interprété comme un futur temporel.

- (7) a. Max **aura** bientôt dix ans.  
b. Non, il a déjà dix ans.

Quand le futur est interprété comme épistémique, il ne peut pas être analysé comme agissant sur le cp. En effet :

(i) il n'est pas sémantiquement constitutif de celui-ci étant donné qu'il ne porte pas sur le prédicat ; ce dernier renvoie à un état de choses qui est en train de se produire ou qui s'est déjà produit au moment de l'énonciation. La prédication faite au futur n'est par conséquent pas réfutable.

- (8) a. Quel âge donnes-tu à Max ?  
b. Il **aura** à peu près dix ans  
??c. Non, il a à peu près dix ans

(ii) Le futur dans sa valeur épistémique n'est pas non plus constitutif du composant attitudinal qui porte sur le cp, comme l'est l'adverbe *peut-être*. En effet, un adverbe comme *peut-être* peut être corrigé par un autre adverbe pour modifier la façon dont est évalué le cp, mais le futur ne peut pas être corrigé dans ce même sens.

- (9) a. Pourquoi est-ce qu'il n'est pas venu ?  
b. Il est **peut-être** fâché  
a. Il est certainement fâché / Il est fâché

- (10) a. Pourquoi est-ce qu'il n'est pas venu ?  
b. Il **sera** fâché  
a. Il est fâché

La rectification faite dans le premier cas (9) peut soit changer l'évaluation épistémique, soit remettre en cause le fait d'en faire une. Dans le deuxième cas (10), la rectification peut uniquement remettre en cause le fait d'introduire une évaluation épistémique au moyen du futur.

Le niveau auquel agit le futur n'est donc pas celui du cp, mais celui de l'énonciation. Il indique que l'énonciation du contenu p est reportée à un moment ultérieur à celui de l'énonciation effective. Il fonctionne ainsi comme une marque oratoire de prétérition du dire, signalant que le dire qui fait l'objet de l'énonciation effective est reporté. Ce fonctionnement est comparable aux marques de prétérition *je ne vous dis pas que p / je vous dirai que p*, qui signalent que l'on ne dit pas ou que l'on dira plus tard

ce que l'on dit présentement. Quand le futur prend une valeur épistémique, ce report du dire est motivé par l'ignorance du locuteur sur l'état de choses que ce dernier décrit dans son énoncé.

Certains adverbes épistémiques sont susceptibles, comme le futur, d'agir à ces deux niveaux. C'est le cas de *peut-être* dans les trois langues romanes. Cet adverbe peut donner des indications sur le contenu en signalant que l'état de choses n'est pas avéré ou est potentiel. Dans ce cas, il a sa valeur épistémique « standard ». Mais il peut aussi donner des indications sur le mode d'énonciation, comme une marque de prétérition, en signalant que l'acte d'énonciation n'est pas avéré ou est potentiel. Dans ce cas, il ne modalise par le cp en le présentant comme désignant un état de choses probable. Dans un emploi comme celui illustré sous (4) il n'est pas question d'évaluer la probabilité d'être italien pour le locuteur. Cette indication « épistémique » se reporte sur l'acte d'énonciation, qui est présenté comme potentiel. Cette figure de prétérition s'actualise dans un contexte dans lequel le locuteur souhaite ne pas tenir compte de cette énonciation pour son argumentation, d'où la valeur concessive que *peut-être* paraît porter dans ce type de configuration (il est en général suivi de *mais*).

Ce double niveau d'analyse est pertinent pour les marques des trois langues romanes. Pour l'illustrer, nous prenons des exemples proposant une paraphrase qui n'est pas censée reproduire le fonctionnement de la forme en question, mais qui sert à expliciter le mode d'action de celle-ci sur le contenu ou sur l'énonciation. Nous prendrons les exemples concernant le français.

Le cas du futur

(11) Elle l'**épousera**

Dans cet exemple le futur fait partie du contenu propositionnel et donne une représentation de l'action comme ayant lieu à un moment ultérieur au moment de l'énonciation effective.

(12) Il **aura** au maximum une vingtaine d'années

Dans cet exemple le futur n'exprime pas une prédiction. Il agit sur le mode d'énonciation, en signalant fictivement que l'énonciation du contenu p est remise à un temps ultérieur à celui de l'énonciation effective. Ce mode d'action peut être explicité par la paraphrase suivante :

'L'énonciation du contenu « x avoir une vingtaine d'années » est remise à plus tard'.

Le cas de *peut-être*

(13) Elle l'épouse **peut-être**

Dans cet exemple, l'adverbe *peut-être* agit sur le contenu propositionnel ; il signale que l'action d'épouser n'est pas avérée.

(14) Il n'a **peut-être** que dix ans, mais il peut comprendre

Dans cet exemple l'adverbe *peut-être* n'agit pas sur le contenu propositionnel. Il ne donne pas une évaluation épistémique concernant l'état de choses « avoir dix ans », mais il agit sur le mode d'énonciation de ce contenu. Il signale fictivement que l'énonciation de ce contenu est potentielle. Ce mode d'action peut être explicité par la paraphrase suivante :

'Il n'est pas certain que l'énonciation du contenu « avoir dix ans » ait lieu'.

Les valeurs que prennent les formes qui agissent au niveau énonciatif, en l'occurrence la valeur épistémique pour le futur et la valeur concessive pour *peut-être*, sont désignées dans notre cadre par le terme « valeurs rhétoriques ». Elles sont considérées comme *semi-conventionnelles*, car elles varient en fonction des circonstances dans lesquelles est énoncé un certain contenu tout en étant attachées à une forme précise. Par exemple le futur italien concessif (qui fonctionne au niveau énonciatif) peut véhiculer la même valeur que *peut-être* concessif en français (qui fonctionne également au niveau énonciatif), mais cette même valeur est issue d'un mode d'action différent sur l'acte d'énonciation : le futur le remet à un temps ultérieur et *peut-être* le présente comme potentiel. La similarité entre le français et l'italien vient du fait que les circonstances dans lesquelles chacun de ces fonctionnements est exploité, à savoir ne pas tenir compte d'une énonciation pour la poursuite d'une argumentation, sont les mêmes pour le futur italien et pour *peut-être* en français. Ces mêmes circonstances ne conviennent pas pour l'usage du futur en français.

Nous utilisons le terme générique de *sens modaux* pour couvrir à la fois les valeurs liées au fonctionnement énonciatif de la forme (comme la valeur épistémique du futur et la valeur concessive du futur italien et de *peut-être*) et celles issues du fonctionnement de la forme au niveau du contenu propositionnel lorsqu'elle intègre le composant attitudinal (comme la valeur épistémique de *peut-être*).

Cette représentation permet de concevoir les sens modaux comme mouvant d'une forme à l'autre au sein d'une même langue ou au gré des langues. La perspective quantitative que nous allons explorer ci-après permettra de dégager des indices de régularités dans la diffusion de ces sens selon les langues.

### **3. La démarche quantitative**

En prenant appui sur une démarche quantitative, on adressera deux questions concernant la diffusion des sens modaux.

#### **Question 1**

La première question concerne les marques qui sont le plus susceptibles de changer de niveau de fonctionnement. Très concrètement, on se demandera pourquoi c'est *peut-être* qui, en français, a acquis un fonctionnement énonciatif lui permettant de véhiculer une valeur rhétorique concessive et

non, par exemple, *probablement*. En effet, l'emploi de *probablement* en lieu et place de *peut-être* dans un exemple comme (2) est très peu vraisemblable :

(15) ??Je suis **probablement** italienne, mais je n'aime pas la pizza

Similairement, dans une perspective comparative, on s'interrogera sur le fait que le futur français paraît nettement plus rarement utilisé pour véhiculer une valeur épistémique (due également à un fonctionnement énonciatif) que ne l'est le futur italien ou, en ce qui concerne le roumain, le présomptif. Nous tâcherons de comprendre si l'acquisition d'un fonctionnement énonciatif permettant à la valeur épistémique du futur ou à la valeur concessive de *peut-être* de s'actualiser est liée à la fréquence que nous pouvons relever pour ces formes, tous types d'emploi confondus, dans les corpus que nous avons interrogés.

## Question 2

La seconde question concerne le différentiel de fréquence entre le futur épistémique français d'un côté et le futur épistémique italien ainsi que le présomptif roumain de l'autre. S'il est confirmé que le futur français est nettement plus rare que le futur italien ou le présomptif roumain pour transmettre une valeur épistémique, alors existe-t-il des formes du français qui sont utilisées en compensation de ce manque concernant cette valeur du futur ? Elle semble être rendue par *devoir* utilisé dans son emploi épistémique en français. Si cette forme comble le vide laissé par les emplois épistémiques que le futur français n'a pas, elle devrait être plus fréquente en français qu'en italien ou en roumain. Nous verrons que ce n'est pas le cas. En étendant l'analyse à la fréquence d'autres adverbes épistémiques, nous chercherons à comprendre si le déséquilibre constaté entre le français et l'italien concernant les emplois épistémiques du futur est dû au fait que certaines langues, dans les corpus que nous avons pris en considération, font moins souvent recours à des indications épistémiques que d'autres.

## 4. Méthode

En ce qui concerne la première question – aptitude d'une forme à acquérir un fonctionnement énonciatif –, nous avons comparé, dans des genres discursifs similaires, la fréquence du futur tous types d'emploi confondus en italien et en français, ainsi que celle du présomptif en roumain et nous avons mis ces fréquences en rapport avec la fréquence des emplois épistémiques de ces formes dans les trois langues. Comme nous avons constaté une corrélation, nous avons vérifié si elle tient pour d'autres formes, en l'occurrence *peut-être*, qui dans les trois langues a des emplois énonciatifs à côté de ses emplois épistémiques. Nous avons comparé dans ce but la fréquence de *peut-être* dans les trois langues tous types d'emplois confondus avec celle des autres adverbes épistémiques pour lesquels nous n'avons pas identifié aussi clairement des emplois énonciatifs. Nous avons vérifié ainsi si la fréquence d'une forme peut être considérée comme jouant un rôle dans l'acquisition d'emplois énonciatifs.

En ce qui concerne la deuxième question – aptitude d’une langue à recourir à des formes véhiculant des valeurs épistémiques –, en partant de l’observation que le futur épistémique est nettement plus fréquent en italien qu’en français, nous avons comparé la fréquence des adverbes épistémiques et de *devoir* épistémique dans les trois langues. Nous avons vérifié ainsi si les différences de fréquences sont toujours en faveur d’une langue ou au détriment d’une autre ou changent au gré des formes.

## 5. Outils et moyens d’analyse

Nous avons utilisé la plateforme BTLC développée à Cologne par Sacha Diwersy, comprenant une série de corpus annotés et étiquetés dans le cadre du projet Presto coordonné par Peter Blumenthal (université de Cologne) et Denis Vigier (ICAR, université de Lyon 2). Cette plateforme permet la recherche d’unités lexicales et grammaticales, associée à la possibilité de définir les traits de l’usage de ces formes au moyen – entre autres – de l’établissement de concordances, de calcul de fréquences relatives, de calculs de co-occurrences et de colligations textuelles (pour une discussion approfondie de cette notion et de ses rapports avec celle de co-occurrence, voir Legallois 2012).

Pour l’italien et le français, nous avons analysé un corpus de presse constitué d’une année d’un quotidien (Le Monde 2008 pour le français, pour un total de 20 410 766 mots et La Stampa 2002 pour l’italien, pour un total de 31 369 484 mots). Nous avons analysé les deux corpus dans leur globalité et nous avons ajouté à ce corpus global trois sous-sections comparables des deux journaux pour les deux langues : la rubrique économique, la rubrique politique (*Interni* dans La Stampa) et la rubrique consacrée à la présentation et critique des livres (*Le Monde des livres* et *l’Inserito tuttolibri*). Elles ont été consultées en vue d’éviter le mieux possible les biais de corpus en vérifiant par exemple si les futurs ne sont pas cantonnés à un type particulier de rubrique. La plateforme a effectué un calcul automatique du nombre d’occurrences des items recherchés ainsi que de leur fréquence relative (par million de mots). Nous avons ensuite procédé à l’analyse manuelle individuelle, en double aveugle, des occurrences ainsi repérées (dans la totalité ou sur un échantillon significatif selon la taille du corpus/sous-corpus pris en considération). Le corpus du roumain est, lui, composé d’une année et demie (janvier 2014 – août 2015) de numéros de l’hebdomadaire *Dilema* pour un total de 3 170 776 mots. Les occurrences recherchées ont été extraites au moyen de Sketch Engine. Pour les emplois énonciatifs de *peut-être*, nous avons tout d’abord retenu les *peut-être* suivi par l’occurrence du mot *mais* dans un co-texte proche (de 0 à 20 mots), car c’est plus particulièrement dans ce type de configuration qu’il peut être utilisé pour signaler qu’une énonciation est potentielle (le locuteur ne la prenant pas en considération dans la suite de son discours – ce qui est indiqué par *mais*). Nous avons ensuite procédé à une analyse des occurrences obtenues, retenant comme énonciatifs seulement les cas dans lesquels l’énoncé où apparaît *peut-être* présente un contenu qui n’est pas susceptible d’être modalisé épistémiquement, mais est considéré comme faisant partie du background discursif selon la définition de concession donnée en Rossari (2014) et Rossari (à paraître). Un tel statut a été testé notamment au moyen de la substitution par des adverbes épistémiques tels que *probablement*, qui n’a pas la possibilité de porter sur une énonciation.

Ci-dessous, les tableaux regroupant les résultats de l’analyse quantitative pour les trois langues.

**TABLEAUX RECAPITULATIFS FREQUENCES BTLC + des fréquences calculées  
manuellement pour le roumain à partir de Sketch Engine**

**T1: Le Monde 2008 / La Stampa 2002 / Dilema Veche, janvier 2014 - août 2015**

	Nombre total d'occurrences de l'item dans le corpus	Nombre total d'emplois concessifs <sup>3</sup>	Fréquence relative de l'item (n. d'occurrences par Mio. de mots)	Fréquence relative emplois concessifs	Taille du corpus en nombre de mots
<i>sûrement</i>	291		14,26		20 410 766
<i>sicuramente</i>	1862		59,36		31 369 484
<i>sigur</i>	199		62,76		3 170 776
<i>probablement</i>	724		35,47		20 410 766
<i>probabilmente</i>	2771		88,33		31 369 484
<i>probabil</i>	466		146,96		3 170 776
<i>sans doute</i>	2483		121,67		20 408 283 <sup>4</sup>
<i>senza dubbio</i>	298		9,50		31 369 186
<i>fara îndoiala</i>	27		8,51		3 170 776
<i>peut-être</i>	2904		142,28		20 410 766
<i>forse</i>	11773		375,30		31 369 484
<i>poate</i>	1051		331,46		3 170 776
<i>peut-être [...] mais</i>	432	14	21,17	0,69	20 406 946
<i>forse [...] ma</i>	1494	63	47,65	2,00	31 354 131
<i>poate [...] dar</i>	40	3	12,61	0,94	3 170 776
		<b>Nombre total d'emplois épistémiques</b>		<b>Fréquence relative emplois épistémiques</b>	
<i>devoir</i>	29327	13 sur 1500	1436,84	12,45	20 410 766
<i>dovere</i>	50396	29 sur 1500	1606,53	31,05	31 369 484
<i>a trebui</i>	2870	52	905,14	16,40	3 170 776
<i>futur</i>	53139	4 sur 1500	2603,48	6,94	20 410 766
<i>futuro</i>	143063	25 sur 1500	4560,58	76,00	31 369 484
<i>présomptif</i>	198	154	62,44	48,56	3 170 776

<sup>3</sup> Pour simplifier la formulation nous parlons d'emplois concessifs pour les adverbes correspondant à *peut-être* alors que nous devrions parler d'emplois énonciatifs chargés d'une valeur concessive. Parallèlement nous parlons d'emplois épistémiques pour le futur alors que nous devrions parler d'emplois énonciatifs chargés d'une valeur épistémique. Enfin pour les formes lexicales (*devoir* ainsi que les adverbes), nous rappelons que pour nous il ne s'agit pas d'emplois énonciatifs. Il s'agit de formes véhiculant directement une valeur épistémique.

<sup>4</sup> La plateforme indique une taille de corpus et de sous-corpus légèrement variable en fonction du nombre d'occurrences de mots quand l'item recherché est composé de plusieurs unités. Elle déduit automatiquement ce nombre de la taille totale du corpus.



T2 : Le Monde 2008, section *Économie* / La Stampa 2002, section *Economia*

	Nombre total d'occurrences de l'item dans le sous-corpus	Nombre total d'emplois concessifs	Fréquence relative de l'item (n. d'occurrences par Mio. de mots)	Fréquence relative emplois concessifs	Taille du sous-corpus en nombre de mots
<i>sûrement</i>	4		6,01		665 119
<i>sicuramente</i>	130		49,78		2 611 269
<i>probablement</i>	38		57,13		665 119
<i>probabilmente</i>	258		98,80		2 611 269
<i>sans doute</i>	57		85,71		665 062
<i>senza dubbio</i>	20		7,66		2 611 249
<i>peut-être</i>	84		126,29		665 119
<i>forse</i>	557		213,31		2 611 269
<i>peut-être [...] mais</i>	13	5	19,55	7,52	665 015
<i>forse [...] ma</i>	89	23	34,09	8,81	2 610 459
		<b>Nombre total d'emplois épistémiques</b>		<b>Fréquence relative emplois épistémiques</b>	
<i>devoir</i>	1194	13	1795,17	19,54	665 119
<i>dovere</i>	5453	22 sur 1000	2088,26	45,94	2 611 269
<b>futur</b>	1724	0 sur 1000	2592,01		665 119
<b>Futuro</b>	16165	3 sur 1000	6190,48	18,57	2 611 269

T3 : Le Monde 2008, section *Politique* / La Stampa 2002, section *Interni*

	Nombre total d'occurrences de l'item dans le sous-corpus	Nombre total d'emplois concessifs	Fréquence relative de l'item (n. d'occurrences par Mio. de mots)	Fréquence relative emplois concessifs	Taille du sous-corpus en nombre de mots
<i>sûrement</i>	20		9,3347		2 142 550
<i>sicuramente</i>	219		64,31		3 405 266
<i>probablement</i>	75		35,00		2 142 550
<i>probabilmente</i>	245		71,95		3 405 266
<i>sans doute</i>	139		64,88		2 142 411
<i>senza dubbio</i>	32		9,40		3 405 234
<i>peut-être</i>	177		82,61		2 142 550
<i>forse</i>	1162		341,24		3 405 266
<i>peut-être [...] mais</i>	33	11	15,40	5,13	2 142 243
<i>forse [...] ma</i>	147	36	43,19	10,58	3 403 745
		<b>Nombre total d'emplois épistémiques</b>		<b>Fréquence relative emplois épistémiques</b>	
<i>devoir</i>	4290	5 sur 1000	2002,29	10,01	2 142 550
<i>dovere</i>	7130	12 sur 1000	2093,81	25,12	3 405 266
<i>futur</i>	6581	1 sur 1000	3071,57	3,07	2 142 550
<i>futuro</i>	16718	7 sur 1000	4909,45	34,37	3 405 266

T4 : Le Monde 2008, section *Monde des livres* / La Stampa 2002, section *Inserito tuttolibri*

	Nombre total d'occurrences de l'item dans le sous-corpus	Nombre total d'emplois concessifs	Fréquence relative de l'item (n. d'occurrences par Mio. de mots)	Fréquence relative emplois concessifs	Taille du sous-corpus en nombre de mots
<i>sûrement</i>	23		20,99		1 095 656
<i>sicuramente</i>	80		56,27		1 421 663
<i>probablement</i>	34		31,03		1 095 656
<i>probabilmente</i>	81		56,97		1 421 663
<i>sans doute</i>	213		194,44		1 095 443
<i>senza dubbio</i>	18		12,66		1 421 645
<i>peut-être</i>	289		263,77		1 095 656
<i>forse</i>	754		530,36		1 421 663
<i>peut-être [...] mais</i>	32	12	29,21	10,95	1 095 403
<i>forse [...] ma</i>	105	33	73,91	23,23	1 420 633
		<b>Nombre total d'emplois épistémiques</b>		<b>Fréquence relative emplois épistémiques</b>	
<i>devoir</i>	690	23	629,76	20,99	1 095 656
<i>dovere</i>	1214	41	853,93	28,84	1 421 663
<i>futur</i>	2162	18	1973,25	16,43	1 095 656
<i>futuro</i>	3384	80	2380,31	56,27	1 421 663

## 6. Analyse des résultats

### 6.1 Eléments de réponse concernant la question 1 (lien entre fréquence de l'emploi d'une forme et sa capacité à acquérir un sens énonciatif)

Le futur italien tous types d'emploi confondus est nettement plus fréquent que le futur français. Il est 1,7 fois plus fréquent dans le corpus global, 2,4 fois dans le sous-corpus économie, 1,6 fois dans le sous-corpus politique et 1,2 fois dans le sous-corpus livres. Cette différence de fréquence est donc constante. A celle-ci est corrélée une différence de fréquence encore plus sensible entre les emplois épistémiques en français et en italien. Le futur dans son emploi épistémique est en effet 10,9 fois plus fréquent en italien qu'en français dans le corpus global, 11,2 fois plus fréquent en italien dans le sous-

corpus politique et 3,4 fois plus fréquent en italien dans le sous-corpus des livres<sup>5</sup>. L'emploi énonciatif du futur semble donc être corrélé à sa fréquence tous types d'emploi confondus. Le recours aux trois langues romanes permet de conforter l'hypothèse d'une corrélation entre fréquence d'emploi (tous types confondus) et acquisition d'emplois énonciatifs. L'adverbe *peut-être* et ses correspondants *forse* et *poate*, qui (selon la démarche qualitative) sont les plus naturels dans une séquence dans laquelle ils ne servent pas à évaluer la probabilité d'un état de choses (fonctionnement énonciatif), arrivent en tête de file concernant leur fréquence tous types d'emploi confondus dans le corpus global ainsi que dans les trois sous-corpus. L'hypothèse d'une corrélation possible entre la fréquence d'une forme tous types d'emploi confondus et sa capacité à acquérir des emplois énonciatifs est consolidée par cette analyse quantitative.

## 6.2 Eléments de réponse concernant la question 2 (y a-t-il des différences ou des ressemblances entre les langues concernant l'exploitation d'indices épistémiques ?)

La fréquence de *devoir* épistémique en français ne permet pas d'envisager une compensation concernant les emplois épistémiques du futur italien. Il est 2,5 fois moins fréquent en français qu'en italien dans le corpus global, 2,3 fois moins fréquent dans le sous-corpus économie, 2,5 fois moins fréquent dans le sous-corpus politique et 1,4 fois moins fréquent dans le sous-corpus livres. La question d'une différence concernant l'exploitation de formes à valeur épistémique selon les langues a donc du sens. En prenant en compte le roumain, on relève que :

(i) La fréquence des emplois épistémiques de *a trebui* (le correspondant de *devoir* épistémique) est 1,9 fois inférieure à celle des emplois épistémiques de *dovere* et 1,3 fois supérieure à celle des emplois épistémiques de *devoir*.

(ii) La fréquence des adverbess épistémiques italien et roumain *poate/forse*, *probabil/probabilmente* et *sigur/sicuramente* est supérieure à celle de leur homologue français. *Peut-être* est 2,3 fois moins fréquent que *poate* en roumain et 2,6 fois moins fréquent que *forse* en italien ; *probablement* est 4,1 fois moins fréquent que *probabil* en roumain et 2,5 fois moins fréquent que *probabilmente* en italien ; enfin *sûrement* est 4,4 fois moins fréquent que *sigur* en roumain et 4,2 fois moins fréquent que *sicuramente* en italien.

(iii) Cette régularité entre français d'un côté et roumain et italien de l'autre est rompue avec l'adverbe *sans doute*. Celui-ci est 12,8 fois plus fréquent que son homologue italien et 14,3 fois plus fréquent son homologue roumain. Dans ces deux langues la fréquence de cet adverbe est très similaire. Ces résultats font ressortir une forme d'homogénéité entre le roumain et l'italien. Dans les corpus consultés, les différences de fréquence sont systématiquement similaires dans ces deux langues par rapport au français : les adverbess les plus fréquents sont les mêmes en italien et en roumain et dépassent nettement la fréquence des adverbess correspondant en français ; vice-versa, *sans doute* est plus fréquent en français que ses homologues roumain et italien dont la fréquence est à nouveau très similaire (8,5 occurrences par million de mots *pour fara îndoiala* et 9,5 pour *senza dubbio*).

(iv) Le roumain paraît se particulariser par rapport aux formes modales de l'italien et du français en ce qui concerne le présomptif. Premièrement, cette forme est nettement plus rare que les autres formes

---

<sup>5</sup> Nous ne prenons pas en compte l'échantillon testé pour le sous-corpus économie, car nous n'avons trouvé aucune occurrence du futur épistémique français.

épistémiques (mis à part *fara îndoiala*, dont la fréquence est faible comme celle de son pendant italien). Deuxièmement, la valeur épistémique du présomptif paraît s'être conventionnalisée, étant donné que 77,8% des occurrences de cette forme véhiculent une telle valeur. Cette proportion est diamétralement différente pour le futur italien (avec seulement 1,7% d'emplois épistémiques pour le futur italien) et a fortiori pour le français (0,3%). Toutefois cette dissemblance est trompeuse. Si on prend en compte la fréquence relative des présomptifs épistémiques dans le corpus du roumain ainsi que celle des futurs italien et français dans une sous-section des deux corpus respectifs (le sous-corpus 'livres'), on relève que cette fréquence est très semblable en italien et en roumain : 56,2 occurrences par million de mots pour le futur épistémique italien et 48,6 pour le présomptif épistémique en roumain, contre 16,4 occurrences par million de mots pour le futur épistémique français dans la sous-section correspondante. Cette différence de fréquence concernant le futur épistémique n'est pas non plus compensée par *devoir* épistémique en français. En roumain, *a trebui* épistémique a une fréquence 1,3 fois plus élevée que *devoir* épistémique français. Mis à part *a trebui* épistémique, qui est sensiblement moins fréquent que *dovere* épistémique italien, il semble donc bien se dégager une homogénéité dans l'exploitation des formes modales entre le roumain et l'italien par rapport au français.

## 7. Retour à l'analyse qualitative

Les données chiffrées qui ressortent de l'analyse quantitative rencontrent sur certains points l'analyse qualitative et permettent d'envisager de nouvelles pistes à explorer qualitativement et quantitativement.

- (i) Dans notre article (Rossari *et al.* à paraître), nous avons relevé que le présomptif roumain et le futur italien ont des fonctionnements modaux semblables et se différencient sensiblement du futur français (notamment sur la valeur concessive que seuls le présomptif et le futur italien peuvent véhiculer). Cette similarité est confirmée par les données quantitatives : la fréquence relative du présomptif avec valeur épistémique est similaire à celle du futur italien avec valeur épistémique et est bien supérieure à celle du futur français épistémique.
- (ii) L'analyse quantitative fait ressortir une autre homogénéité concernant le roumain et l'italien : les adverbes épistémiques ont des fréquences semblables en roumain et en italien. L'homogénéité relevée dans l'analyse qualitative concernant le futur et le présomptif semble donc s'étendre sur d'autres formes modales, ce qui laisse présager une exploitation similaire des formes épistémiques dans les deux langues. Cette piste mériterait d'être explorée au moyen d'une démarche qualitative concernant les différentes valeurs que ces adverbes peuvent véhiculer et au plan quantitatif en faisant varier les corpus et les tranches diachroniques qu'ils représentent.

## 8. Conclusions

Partant d'une analyse qualitative selon laquelle :

- (i) il y a plus de proximité entre l'italien et le roumain qu'avec le français en ce qui concerne le déploiement du sens modal de certaines formes (les emplois épistémiques du futur italien et du présomptif sont plus fréquents qu'en français ; le roumain et l'italien peuvent utiliser le futur dans des circonstances où, en français, seul *peut-être* peut être utilisé, cf. emploi concessif du futur) ;
- (ii) certains de ces sens se manifestent quand la forme a un fonctionnement énonciatif,

notre analyse quantitative (a) a confirmé en partie cette proximité, (b) a fait ressortir d'autres points de convergence entre roumain et italien d'une part et français d'autre part et (c) a mis au jour de nouvelles pistes qui méritent à la fois une analyse qualitative et une nouvelle analyse quantitative plus poussée.

- (a) La valeur épistémique portée par les marques de flexion temporelle (futur et présomptif) est nettement supérieure en termes de fréquence en roumain et en italien par rapport au français.
- (b) Les formes lexicales véhiculant une valeur épistémique ont une fréquence semblable en roumain et en italien par rapport au français.
- (c) Les formes les plus fréquentes sont celles qui semblent le plus enclines à avoir un fonctionnement énonciatif.

## **Bibliographie**

- Baranzini Laura (à paraître), « Existe-t-il un futur concessif en italien ? », *Idées de Langue*, Laboratoire LDI (Lexiques, Dictionnaires, Informatique) CNRS UMR 7187.
- Berretta Monica (1997) « Sul futuro concessivo: riflessioni su un caso (dubbio) di de/grammaticalizzazione », *Linguistica e Filologia*, 5, p. 7-41.
- Bertinetto Pier Marco (1991), « Il verbo », in *Grande grammatica italiana di consultazione*, Lorenzo Renzi et Gianpaolo Salvi (éds), Bologna, il Mulino, p. 13-161.
- Dendale Patrick (2001), « Le futur conjectural versus devoir épistémique : différences de valeur et restrictions d'emploi », *Le Français Moderne*, 69, 1, p.1-20.
- Legallois Dominique (2012), « La colligation : autre nom de la collocation grammaticale ou autre logique de la relation mutuelle entre syntaxe et sémantique ? » *Corpus* [en ligne] 11, <http://corpus.revues.org/2202>.
- Morency Patrick (2010), « Enrichissement pragmatique du futur », *Cahier Chronos* 21, p. 197-214.
- Popescu Cecilia M. (2013), *Viitorul si conditionalul în limbile romanice*, Craiova, Editura Universității.
- Rossari Corinne (à paraître), « La présupposition dans les structures concessives », in *La Présupposition entre théorisation et mise en discours*, Amir Biglari, et Marc Bonhomme (éds.), Paris, Classiques Garnier.
- Rossari Corinne (2014), « How does a concessive value emerge ? » In *Pragmatic markers from Latin to Romance languages. Studies in diachronic and historical linguistics*, Chiara Ghezzi et Piera Molinelli (éds.), Oxford, Oxford University Press, p. 237-260.
- Rossari Corinne, Ricci Claudia et Siminiciuc Elena (à paraître), « Les valeurs du futur modal en français, italien et roumain », in *Le futur dans les langues romanes*, Laura Baranzini, Juan Pedro Sanchez-Mendez et Louis De Saussure (éds.) Bern, Peter Lang.